

Statistical Models for the Support of Forensic Fingerprint Identifications

by

Joshua Abraham

A Thesis Submitted for the Degree of
Doctor of Philosophy



Centre for Forensic Science

21st May 2017

CERTIFICATE OF ORIGINAL AUTHORSHIP

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as part of the collaborative doctoral degree and/or fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

I certify that I have received ethics clearance from the appropriate authorities in accordance with UTS policies on human and animal research.

Student's Signature

Date

This research is supported by an Australian Government Research Training Program Scholarship.

Acknowledgements

I would like to thank my supervisors Prof. Claude Roux, Prof. Chris Lennard, and Prof. Christophe Champod, for their guidance and support in pursuing the research topic that I am most passionate about. I would like to also my brother, Daniel, for his motivation, belief and pride in my work. Finally, I would like to thank my beloved parents for their unconditional love, support and patience.

Contents

1	Introduction	1
1.1	History of Fingerprints	2
1.2	Fingerprint Identification	9
1.2.1	The Scientific Basis of Fingerprint Identification	9
1.2.2	Automated Fingerprint Identification Systems (AFISs)	11
1.2.3	ACE-V Identification Methodology	18
1.2.4	Forensic Fingerprint Identification Standards	21
1.3	Criticisms of Fingerprint Identification	24
1.3.1	Admissibility of Scientific Expert Testimony	25
1.3.2	Erroneous Identifications	25
1.3.3	Fingerprint Expert Accuracy under ACE-V	28
1.3.4	ACE-V and Cognitive Bias	28
1.3.5	Individualisation, Uniqueness and Discriminability	30
1.4	Statistical Models and Identification	32
1.4.1	The Need for Statistical Models	32
1.4.2	Identification Conclusions and Statistical Models	33
2	A Review of Statistical Models for Fingerprint Identification	36
2.1	Foundations for Probability of Random Correspondence Models	36
2.2	A Review of Historical Probabilistic Models for Fingerprint Rarity	38
2.2.1	Square Ridge Region Analysis Based Models	38
2.2.2	Other Square Ridge Region Analysis-Based Model Variants	40
2.2.3	Minutiae Even-Based Models	46
2.2.4	Landmark Referencing Models	50
2.3	A Review of Modern Probabilistic Models for Fingerprint Rarity	57
2.3.1	Spatial Homogeneity Probability Models	57
2.3.2	Spatio-Directional Based Generative Models	63
2.3.3	Bayesian Network Based Generative Model	71
2.3.4	Inhomogeneous Spatial Point Process Based Models	73
2.4	Foundations of Likelihood Ratio Models	80
2.5	The Probabilistic Relationship Between PRC and LR models	83
2.5.1	Feature Vector Based LR Models	85
2.6	AFIS Score Based LR Models	95

2.6.1	The Relationship Between AFIS and LR	95
2.6.2	Egli et al. 2006	97
2.6.3	Choi et al. 2011	97
2.6.4	Model Methodology Analysis	97
2.7	Likelihood Ratio Model Assessment	98
2.7.1	Evaluation of Likelihood Ratio Accuracy	99
2.7.2	Empirical Cross Entropy and Calibration of Likelihood Ratios . . .	100
3	An AFIS Candidate List Centric Likelihood Ratio Model Based on the Spatial Analyses of Minutiae	103
3.1	Considerations for Model Development	105
3.1.1	Operational Scenario: AFIS candidate list exemplar-to-fingermark identification	105
3.1.2	General Properties of the Data: Feature Dimensionality, Completeness, and Quality	107
3.1.3	Intrinsic and Extrinsic Factors of Spatial Variability of Landmark Based Features	108
3.2	Feature Vector Background	111
3.2.1	Feature Vector: General Spatial Components	112
3.2.2	Feature Vector Pairwise Spatial Analytical Components	113
3.2.3	Euclidean Distance Matrix Analysis	115
3.2.4	Partial Procrustes Method	116
3.2.5	Thin Plate Spline and Derived Measures	117
3.2.6	Three Dimensional Kolmogorov-Smirnov Statistic for Landmarks . .	120
3.3	Proposed Model	122
3.3.1	Feature Vector Definition	122
3.3.2	Learning Feature Vector Classes with Support Vector Machines . . .	125
3.3.3	Likelihood Ratio Calculations	126
3.3.4	Scalability of Likelihood Ratios	129
3.4	Experimentation	132
3.4.1	Experimental Datasets	133
3.4.2	Experimental Methodology	138
3.4.3	Experiment A: Proof-of-Concept	142
3.4.4	Experiment B: Large Scale Experiment I	149
3.4.5	Experiment C: Large Scale Experiment II	156
3.4.6	Real-World Validation Experiment	179
3.5	Conclusions	182
4	An AFIS Candidate List Centric Likelihood Ratio Model using a Kernel Density Estimator Based Framework	185
4.1	Proposed Model	186
4.1.1	Kernel Density Estimation	186
4.1.2	Feature Vector and Dimensionality	187

4.1.3	Likelihood Ratio Calculations	188
4.2	Experimentation	190
4.2.1	Experiment: KDE Large Scale Experiment with Feature Vector Sub- set Selection Dimensionality Reduction	191
4.2.2	Experiment: KDE Large Scale Experiment with PCA Feature Vec- tor Dimensionality Reduction	194
4.2.3	Real-World Validation Experiment	200
4.3	Conclusions	201
5	A Person-of-Interest Likelihood Ratio Model based on Inter-Minutia Distances	204
5.1	Considerations for Model Development	205
5.1.1	Operational Scenario: POI Model Evaluation for ACE-V Inconclu- sive Results and Quality Assurance Processes	205
5.1.2	Model Applicability and Data Collection Methodology	207
5.2	Proposed Model	207
5.2.1	Within-Source Feature Vector	207
5.2.2	POI Model Within Source Kernel Density Estimators	211
5.2.3	Likelihood Ratio	212
5.3	Experimentation	213
5.3.1	Experiment A: Single Crime Mark	213
5.3.2	Experiment B: Multiple Crime Marks	214
5.4	Conclusions	216
6	General Discussion and Conclusions	218
6.1	Review of Research Objectives	218
6.1.1	Review of Proposed AFIS-Centric Models	219
6.1.2	Review of Proposed Person-of-Interest Model	219
6.1.3	Research Relevance for the Forensic Science Community	220
6.1.4	Future Work	220
	Appendices	222
	A Fingerprint Features	223
	B Information Theory	225

List of Figures

1.1	Ancient Greek pottery seals found in (Waldstein, 1902).	3
1.2	The 9 fingerprint classes that Purkyně defined (images adapted from Cummins et al. (1940)). These are: (a) transverse curves (i.e. simple arch), (b) the central longitudinal stria (i.e. tented arch), (c) the oblique stripe (i.e., loop, ulnar, or radial), (d) the oblique loop (i.e., loop, ulnar, or radial), (e) the almond (i.e., whorl variant), (f) the spiral (i.e., whorl variant), (g) the ellipse (i.e., whorl variant), (h) the circle (i.e., whorl variant), (i) the double whorl.	5
1.3	Herschel’s right index and middle fingers impressed at 1859, 1877, and 1916. Figure sourced from Herschel (1916).	7
1.4	(a) The line that intersects the delta and the centre of the system of arcs (i.e., roughly the core location) is suggested to be used to align fingerprints in the patent proposed by Maurer. (b) Suggested defined regions that can be analysed for matching and classification purposes after alignment is performed.	12
1.5	Flowchart of modern ACE-V process used in conjunction with AFIS. The iterative comparison of each exemplar fingerprint in the AFIS candidate list (i.e., most similar records in an AFIS database ranked by order of similarity) is performed until identification occurs or no more exemplars are left. The red flow lines indicate the process for the verification stage analysis. The purple flow line from the ‘agreement of features’ test shows the ACE process that skips the evaluation stage.	21
1.6	The contentious 16 points used to identify the latent mark retrieved from the Marion Ross murder crime scene (top) against Shirley McKie’s template mark (below). Images sourced from German (2015).	26
1.7	(left) The latent mark retrieved from the crime scene with 10 minutiae (out of the 15 marked features) used for the erroneous identification. (centre) Ouhmane Daoud’s exemplar correctly identified by Spanish Police. (right) Brandon Mayfield’s exemplar erroneously identified by the FBI (OIG., 2006, pp. 132-134).	27
1.8	(left) Brandon Mayfield’s exemplar erroneously identified by the FBI with 5 minutiae not existing in the latent mark. (right) The latent mark retrieved from the crime scene with the 5 missing minutiae positions overlaid (OIG., 2006, pp. 140-141).	29

2.1	(left) Galton squares on two regions of a fingerprint with the covered squares and guessed ridge details. (right) fingerprint with region details transparent. The green square indicates that the region can be guessed correctly from the surrounding regions, whereas the red square cannot. . . .	39
2.2	Sample fingerprint divided into 1 x 1 mm square regions. Image sourced and adapted from Osterburg et al. (1977).	43
2.3	(left) Fingermark from homicide case recorded by Henry (1900) (right) Features retrieved for calculations of latent mark PRC. Images can be found in (Henry, 1900, pp. 53).	46
2.4	(left) Concentric circle structure with origin at the core. (right) Minutia code defining a unique order, derived from the concentric circle structure. .	51
2.5	Stoney's method for finding the focal minutia and neighbours.	53
2.6	Minutia pair spatial and directional match tolerance illustrated on a given fingerprint region. (image adapted from Pankanti et al. (2001)).	60
2.7	(a) A NIST4 (Watson, 1992) fingerprint with categorised minutiae clusters derived from the mixture model with parameters tuned by the EM algorithm. (b) Directional density of mixture model clusters. (c) Spatial density map of mixture model. (d) Three dimensional view of the spatial density map. (e) Three dimensional view of spatial-directional clusters.	65
2.8	(left) Minutiae sequencing initial procedure starting off with locating the closest minutiae to the core point. (right) The resulting Bayesian network representing the minutia dependency sequence.	72
2.9	(left) An over-dispersed (or uniform, regular) point pattern. centre: A random (i.e., CSR) point pattern. (right) A clustered point pattern. . . .	74
2.10	The different sectors (left) and ridge intervals (right) used to define fingerprint regions. Images sourced with permission from Champod (1996). . .	74
2.11	Density maps for different types of fingerprints. Clearly regions near singularities contain higher density in comparison to periphery regions. Hence, density maps clearly reveal the classification of the fingerprints. Images sourced with permission from Champod (1996).	75
2.12	Two different configurations with a similar amount of minutiae but vastly different PRC calculation. This suggests that the numeric standard concept is inadequate. Images sourced with permission from Champod (1996). . . .	76
2.13	AFIS distribution: likelihood ratio for a hypothetical AFIS system. In this example, the score of 52 will favour the hypothesis in support of the defence (i.e., x and y were produced by different fingers).	82
2.14	The Delaunay triangulation (left) and radial triangulation (right) differences for a configuration of 7 minutiae. The blue point for the radial triangulation illustration represents the centroid (i.e., arithmetic mean of minutiae x-y coordinates).	85

2.15	Radial triangulation structures for the corresponding configurations of minutiae from the same finger source. Due to distortion, the centroid structure has different triangles for the lower triangles that will lead to erroneous results for the models by Neumann et al. (2012) and Neumann et al. (2015).	94
2.16	Doddington's Zoo Plot- scatterplot of average genuine (within-finger) versus average imposter (between-finger) AFIS scores for candidate list entries of each fingerprint template found in a hypothetical population. If the AFIS scores are correlated or are a monotonically increasing function of the probabilities used in a LR model, template sub-populations of categories Chameleons, Phantoms, and Goats are more likely to produce LR values incorrectly supporting the hypothesis in favour of the defence when comparing features sourced from within-finger query samples. In addition, template sub-populations of Worms, Goats, and Wolves/Lambs are more likely to produce LR values incorrectly supporting the hypothesis in favour of the prosecution when comparing features sourced from between-finger query samples. Sub-populations of Doves/Sheep are more likely to produce LR values in favour of the correct hypothesis.	96
2.17	Tippett plots showing the cumulative distributions for LRs where H_P and H_D cases are true (blue and red lines, respectively). The false H_P and H_D proportion of cases for a decision threshold value of $\log_{10}(LR) = 1$ is depicted.	100
2.18	ECE plots showing the ECE value against a range of log odds ratio values (with each plot showing results for observed LR values (red), neutral LR values of $LR=1$ (black), and calibrated LRs (blue)). The ECE plots show cases where (top left) uncalibrated LR values are better for prediction than no information (i.e., $LR=1$ always), (top right) uncalibrated LR values being worse than the neutral constant of $LR=1$ for log odds ratios approximately greater than 0.75, (bottom left) uncalibrated LRs having perfect discrimination between H_P and H_D cases, and (bottom right) a perfectly discriminated LR set with observed values having extremely poor calibration (for the given log prior odds range).	102
3.1	Flowchart illustrating how the model can be incorporated with AFIS and the ACE portion of the ACE-V methodology. This includes re-ordering the candidate list by the LR on unadjusted automated correspondences found between the candidate and crime mark, filtering out candidate list entries with LR values less than a defined threshold, and incorporating a re-calculated LR value on expert adjusted correspondences from the Comparison stage to contribute as additional analytical evidence (along with expert markup notes and analysis) for the Evaluation stage.	106

3.2	a) A configuration of 5 minutiae submitted for a search on an AFIS. b) Spatial distribution of Matches (blue) and Close Non-matches (Red) for the given search configuration. c) Spatial distribution for a specific minutia of the match (i.e., within-source) population that has bimodal peaks and is non-Gaussian.	109
3.3	(top left) Fingerprint with a marked configuration of minutiae used for an AFIS search. (top right) A corresponding configuration of minutiae from a rolled fingerprint from the same source finger. (bottom) The polygons created from both minutiae correspondences centred at the origin (0,0) (marked by the black point) and aligned using the partial Procrustes method. The respective geometric medians are represented by the red and green '+' symbols.	114
3.4	(a) Query configuration of 7 minutiae. (b) A corresponding configuration of minutiae from a different finger (i.e., simulated close non-match). The corresponding inter-minutia distances used to find the largest and smallest distance ratios found in the form distance matrix to be used by the EDMA test statistic are illustrated by green and red lines, respectively, on both configurations.	116
3.5	(a) Query configuration of 7 minutiae. (b) A corresponding configuration of minutiae from a different finger (i.e., close non-match). (c) The partial Procrustes method of alignment applied.	117
3.6	(a) Query configuration of minutiae and corresponding configuration from a different finger. The corresponding minutiae are illustrated between the two fingerprints. (b) The respective TPS deformation grid found without prior alignment of configurations. In the proposed method, the partial Procrustes method is applied beforehand, resulting in much smaller affine transformation effects (such as rotation) on the grid. However, this example without prior alignment is used to clearly illustrate the affine component of TPS.	120
3.7	(a) The region point tallies (location only) of the 2D K-S statistic illustrated for an arbitrary minutia from the minutiae query configuration found in Figure 3.4 (a). (b) The respective region tally for the corresponding minutia from the close non-match configuration found in Figure 3.4 (b).	122
3.8	SVM framework for calculating LR on defined feature vectors.	129
3.9	Random selection of samples in Dataset A for a particular finger where various direction, torsion and pressure applications are observed.	134
3.10	Random selection of samples in Dataset C for a particular finger with different various directional application.	135
3.11	(left) A configuration of minutiae from a latent mark. (right) Ten-print exemplar from the NIST27 database with corresponding marked minutiae.	136

3.12	(left) Aligned corresponding minutiae from the NIST27 sample illustrated in Figure 3.11. (right) Thin plate spline deformation grid illustrating the encountered distortion using the corresponding minutiae pairs. This dataset is an ideal evaluation set as it will contain real life spatial variability introduced from skin distortion, the fingerprint expert in the precision of his/her markings and other the crime scene environmental factors.	137
3.13	(left) Marked configuration of minutiae from a fingerprint. (right) A paired corresponding close non-match configuration of minutiae (marked in yellow) from another fingermark.	137
3.14	Overview of the minutiae correspondence search algorithm. The partial procrustes and TPS related constraints are highlighted in blue and pink areas, respectively.	139
3.15	Flowcharts detailing procedure for the different alignment stages in the minutiae correspondence search algorithm.	140
3.16	Screenshot of the simple fingerprint editor tool used to edit/remove automatically detected minutiae and add additional minutiae. The example displayed has numerous spurious low-quality minutiae that need to be removed.	142
3.17	Tippett plots for configurations with 5 (top left), 6 (top right), 7 (bottom left), and 8 (bottom right) minutiae. The x -axes represents the logarithm (base 10) of the LR_K values in equation (3.2) for match (blue line) and close non-match (red line) populations, while the y -axes represents proportion of such values being greater than x . The green vertical dotted line at $x = 0$ signifies a marker for $LR_K = 1$. The non-match LR_K value gaps are due to the small sample sizes and the random selections used in the two-folds cross-validation training, producing dissimilarly fitted sigmoid functions on respective training sets.	145
3.18	A real life close non-match sourced from www.clpex.com . The AFIS score (3M Cogent) was reported as 1135. The LR_K value is relatively low at $\log(0.023) = -1.64$	146
3.19	A close non-match found in Langenburg (2009). The respective $\log LR_K$ value is $\log(0.0297) = -1.53$	147
3.20	A real life close non-match (top) sourced from www.clpex.com . The AFIS score (NEC) was reported as approximately 3800. The LR_K value range for $\sum_{i=5}^8 \binom{9}{m} = 255$ sub configurations containing $m = 5, 6, 7$, and 8 minutiae are illustrated on the Tippett plots (bottom) in respective order.	147
3.21	A true correspondence found in the within-finger distortion set. This configuration was not used for training or evaluation. The respective $\log LR_K$ value is $\log(213.1) = 2.33$	148
3.22	A true correspondence with 7 minutiae found in the within-finger distortion set. This configuration of minutiae was not used for training or evaluation. The respective $\log LR_K$ value was $\log(29.71) = 1.47$	148

3.23	A true correspondence found in the within-finger distortion set. The respective $\log LR_K$ value was $\log(0.041) = -1.39$	148
3.24	Box plot detailing the distribution of the number of close non-matches for sub configurations containing $m = 4, 5$, and 6 minutiae (top) and $m = 7, 8$, and 9 minutiae (bottom). It is evident that an increase in the number of corresponding minutiae generally results in less close non-matches occurring.	150
3.25	The distribution of centroid for search configurations used in the experiment. The core location for the respective finger for each search configuration is at the origin, (0,0). It is clear that the left periphery regions are favoured. This is due to the sampling method, where searching is starting from selecting k-configurations of minutiae using the left most minutiae as an initial reference point, while the large number of possible configurations along with the computational complexity of searching did not allow the experiment to completely exhaust all cases of k-configurations amongst the right periphery of fingerprints.	151
3.26	A plot of the average Ripley's K function values taken at 15 intervals for configurations with centroids found at different regions of the fingerprint. In agreement with result found in Chen et al. (2008) (see Section 2.3.4.2), the minutiae point patterns have a tendency to be over-dispersed in the micro scale, but tend to cluster from measure of 25 pixels and above (for 500ppi images of fingerprints).	152
3.27	The distribution of calibrated LR_K values of close non-matches and matches from models created for configurations with 4 to 9 minutiae.	154
3.28	The distribution of calibrated LR_{weight} values of close non-matches and matches from models created for configurations with 4 to 9 minutiae.	155
3.29	The LR_{weight} ECE plots for 4 to 9 minutiae models. The non-calibrated values are worse to use than having no information, suggesting that calibration is required. The calibrated values show a trend of increased dichotomy between close non-matches and matches from 4 to 7 minutiae models. The performance anomalies of 8 and 9 minutiae models are evident, as the trend is not continued.	155
3.30	EDMA statistic distributions for randomly sampled corresponding configurations (up to 100,000 match and close non-matches) for configurations with 4 (top left), 5 (top middle), 6 (top right), 7 (bottom left), 8 (bottom middle), and 9 (bottom right) minutiae.	157
3.31	TPS bending energy measure distributions for randomly sampled corresponding configurations (up to 100,000 match and close non-matches) for configurations with 4 (top left), 5 (top middle), 6 (top right), 7 (bottom left), 8 (bottom middle), and 9 (bottom right) minutiae.	158

3.32	TPS angle distributions for randomly sampled corresponding configurations (up to 100,000 match and close non-matches) for configurations with 4 (top left), 5 (top middle), 6 (top right), 7 (bottom left), 8 (bottom middle), and 9 (bottom right) minutiae.	158
3.33	TPS shear metric distributions for randomly sampled corresponding configurations (up to 100,000 match and close non-matches) for configurations with 4 (top left), 5 (top middle), 6 (top right), 7 (bottom left), 8 (bottom middle), and 9 (bottom right) minutiae.	159
3.34	TPS scale metric distributions for randomly sampled corresponding configurations (up to 100,000 match and close non-matches) for configurations with 4 (top left), 5 (top middle), 6 (top right), 7 (bottom left), 8 (bottom middle), and 9 (bottom right) minutiae.	159
3.35	TPS offset metric distributions for randomly sampled corresponding configurations (up to 100,000 match and close non-matches) for configurations with 4 (top left), 5 (top middle), 6 (top right), 7 (bottom left), 8 (bottom middle), and 9 (bottom right) minutiae.	160
3.36	The centroid size difference metric distributions for randomly sampled corresponding configurations (10,000 match and close non-matches) for configurations with 4 (top left), 5 (top middle), 6 (top right), 7 (bottom left), 8 (bottom middle), and 9 (bottom right) minutiae.	161
3.37	Ordinary Sum of Squares distributions for randomly sampled corresponding configurations (10,000 match and close non-matches) for configurations with 4 (top left), 5 (top middle), 6 (top right), 7 (bottom left), 8 (bottom middle), and 9 (bottom right) minutiae.	161
3.38	KS statistic distributions for randomly sampled corresponding configurations (10,000 match and close non-matches) for configurations with 4 (top left), 5 (top middle), 6 (top right), 7 (bottom left), 8 (bottom middle), and 9 (bottom right) minutiae.	162
3.39	Geometric median difference measure distributions for randomly sampled corresponding configurations (10,000 match and close non-matches) for configurations with 4 (top left), 5 (top middle), 6 (top right), 7 (bottom left), 8 (bottom middle), and 9 (bottom right) minutiae.	162
3.40	Polygon area difference measure distributions for randomly sampled corresponding configurations (10,000 match and close non-matches) for configurations with 4 (top left), 5 (top middle), 6 (top right), 7 (bottom left), 8 (bottom middle), and 9 (bottom right) minutiae.	163
3.41	Correlations of feature vector components of match corresponding configurations.	164
3.42	Correlations of feature vector components of close-non match corresponding configurations.	165

3.43	Principal Component Analysis plot for the first three principal components (x-axis PC1, y-axis PC2, z-axis PC3) for corresponding configurations (2,000 match and close non-match random samples plotted in blue and red, respectively) for configurations with 4 (top left), 5 (top middle), 6 (top right), 7 (bottom left), 8 (bottom middle), and 9 (bottom right) minutiae. All of the PCA plots suggest that the close non-match and match populations are not linearly separable through the principle components as there is substantially overlay between both groups, making them indistinguishable.	166
3.44	Scree plots of the analytical portion of the feature vectors for match (left) and close non-match (right) configurations. Match configurations of sizes 4, 5, and 6 and close non-match configurations of sizes 4, 5, 6, 7, 8 and 9 have multiple points of inflexion.	167
3.45	Variable factor map (PCA) for the match population, depicting a view of the projection of the variables projected into the plane spanned by the first two principal components.	168
3.46	Variable factor map (PCA) for the close non-match population, depicting a view of the projection of the variables projected into the plane spanned by the first two principal components.	169
3.47	(top left) First two principal components of feature vectors derived from a random sample of 2000 match and close non-match examples (total 4000) for configurations of 8 minutiae. This is followed by the illustrated classification of the first two principal components of match and close non-match examples using several different methods with two-folds cross-validation. The accuracy of each method is reported in the lower right corner of each sub-plot as the proportion of successful classifications.	170
3.48	The distribution of $\log_{10}(LR_K)$ for configurations with 4 (top left), 5 (top middle), 6 (top right), 7 (bottom left), 8 (bottom middle), and 9 (bottom right) minutiae. Generally speaking, an increased accuracy is observed with increased number of minutiae.	172
3.49	The distribution of $\log_{10}(LR_{weight})$ for configurations with 4 (top left), 5 (top middle), 6 (top right), 7 (bottom left), 8 (bottom middle), and 9 (bottom right) minutiae. Generally speaking, an increased accuracy is observed with increased number of minutiae.	172
3.50	The LR_{weight} ECE plots for 4 to 9 minutiae models. The results are very similar to those from the previous experiment.	173
3.51	The sigmoid functions fitted on raw SVM output values of models from which posterior probabilities are calculated for configurations with 4 (top left), 5 (top middle), 6 (top right), 7 (bottom left), 8 (bottom middle), and 9 (bottom right) minutiae.	174

3.52	Cumulative Match Discriminatory Analysis for feature vectors (mandatory components) from configuration correspondences with 4 (top left), 5 (top middle), 6 (top right), 7 (bottom left), 8 (bottom middle), and 9 (bottom right) minutiae. Generally speaking, an increased accuracy is observed with increased number of minutiae.	176
3.53	Cumulative Match Discriminatory Analysis for feature vectors (all components) from configuration correspondences with 4 (top left), 5 (top middle), 6 (top right), 7 (bottom left), 8 (bottom middle), and 9 (bottom right) minutiae. Generally speaking, an increased accuracy is observed with increased number of minutiae.	177
3.54	Thin Plate Spline (TPS) distortion grid for the (left) incorrect suspect Mayfield and (right) correct suspect Daoud.	181
3.55	Box plot of the LR_K values for 120 sub-sampled correspondences of 7 minutiae for the Madrid bombing correspondences from Mayfield (incorrect) and Daoud (correct).	181
4.1	KDE-based model framework for calculating LR on defined feature vectors.	190
4.2	The distribution of calibrated LR_K values of close non-matches and matches for models created for configurations of 4 to 9 minutiae.	193
4.3	The distribution of calibrated LR_{weight} values of close non-matches and matches for models created for configurations of 4 to 9 minutiae.	194
4.4	The distribution of $\log_{10}(LR_K)$ for configurations with 4 (top left), 5 (top middle), 6 (top right), 7 (bottom left), 8 (bottom middle), and 9 (bottom right) minutiae.	197
4.5	The distribution of $\log_{10}(LR_{weight})$ for configurations with 4 (top left), 5 (top middle), 6 (top right), 7 (bottom left), 8 (bottom middle), and 9 (bottom right) minutiae.	197
4.6	Cumulative Match Characteristic plots of the proposed KDE-based method for configuration correspondences with 4 (top left), 5 (top middle), 6 (top right), 7 (bottom left), 8 (bottom middle), and 9 (bottom right) minutiae. A strict increase in rank-1 identification accuracy is observed with an increase in the number of minutiae.	198
5.1	Flowchart illustrating how the POI model may be used in practice following an ACE-V identification evaluation on existing fingerprint records of a POI and the crime mark(s). If an inconclusive decision is made in the evaluation or a quality assurance step is required, the POI model can be built and used to decide if further human expert analysis is required to verify or resolve issues with the original identification decision.	206
5.2	A visual representation of the form matrix feature vectors for the same configuration of minutiae sourced from two impressions of the same finger. The colour map is used to indicate the measure of scaled inter-minutia distances.	209

5.3	The feature vector for a given configuration of minutiae under different force and directional impression applications. While small difference are introduced to the structure from different distortion, the structure remains largely similar.	211
5.4	The POI LR model numerator calculation framework.	212
5.5	(left) box plots of the LR values for matches with 4 to 9 corresponding minutiae. (right) box plots of the LR values for close non-matches with 4 to 9 corresponding minutiae.	214
5.6	(left) box plots of the average LR values of two singularly sourced crime marks that are true matches containing 4 to 9 corresponding minutiae. (right) box plots of the average LR values of two singularly sourced crime marks that are close non-matches containing 4 to 9 corresponding minutiae.	215
5.7	(left) box plots of the average LR values of five singularly sourced crime marks that are true matches containing 4 to 9 corresponding minutiae. (right) box plots of the average LR values of five singularly sourced crime marks that are close non-matches containing 4 to 9 corresponding minutiae.	216
A.1	(top) Most common fingerprint pattern classifications as defined in Henry (1900). (bottom left) The most common level 2 features (bifurcation and ridge ending minutiae). (bottom right) Level 3 features of open/closed pores and local ridge detail.	223
A.2	Real world latent marks sourced from the NIST27 database (see Garriss et al. (2000)).	224

List of Tables

2.1	Osterburg’s relative frequencies for defined characteristics.	45
2.2	Kingston’s relative frequencies of 2464 minutiae (Kingston, 1964).	49
2.3	Quality map definition used in Roxburgh (1934).	51
2.4	Sample correspondence probability calculations for the model reported by Pankanti et al. (2001).	61
2.5	Ridge and non-ridge PRC values from the FVC2002 DB1 database	70
2.6	Some likelihood ratio error rate results for different finger/region combinations.	87
3.1	Details on the between-finger datasets used in the experiments.	135
3.2	The discovered corresponding configuration of minutiae statistics from searches on Dataset A (match cases) and Dataset D (close non-match cases).	143
3.3	The LR_K rates of misleading evidence in favour of defence (RMED) and prosecution (RMEP) for corresponding configurations of $n = 5 \dots 8$ minutiae found in the evaluation set derived from minutiae correspondence searches in Dataset A.	144
3.4	The discovered statistics of corresponding configurations of $n = 4 \dots 9$ minutiae resulting from searches on Datasets B, C (match cases) and Dataset E (close non-match cases).	149
3.5	The LR_K and LR_{weight} rates of misleading evidence in favour of defence (RMED) and prosecution (RMEP) for corresponding configurations of $n = 4 \dots 9$ minutiae found in the evaluation set derived from minutiae search configurations in Dataset B.	153
3.6	The LR_K rates of misleading evidence in favour of defence (RMED) and prosecution (RMEP) for corresponding configurations of $n = 4 \dots 9$ minutiae found in the evaluation set for Experiments B and C.	171
3.7	The LR_{weight} rates of misleading evidence in favour of defence (RMED) and prosecution (RMEP) for corresponding configurations of $n = 4 \dots 9$ minutiae found in the evaluation set for Experiments B and C.	171
3.8	The LR_K and LR_{weight} rates of misleading evidence in favour of defence (RMED) and prosecution (RMEP) of the corresponding configurations of $n = 4 \dots 8$ minutiae found in the evaluation set for models created per region/minutiae size.	175

3.9	The $RMEP_R$ and $RMED_R$ of each model (for $n = 4 \dots 6$ minutiae) for the corresponding configurations in the evaluation set resulting from different gallery sizes.	178
3.10	The $RMEP_R$ and $RMED_R$ of each model (for $n = 7 \dots 9$ minutiae) for the corresponding configurations in the evaluation set resulting from different gallery sizes.	179
3.11	The LR_K rates of misleading evidence in favour of defence (RMED) for sub-sampled corresponding configurations of $n = 4 \dots 7$ found in the NIST27 validation set (Dataset F) containing only true correspondences.	180
3.12	The LR_K rates of misleading evidence in favour of prosecution (RMEP) for sub-sampled corresponding configurations of $n = 4 \dots 7$ minutiae found in the Dataset G Close Non-Match validation set containing only true correspondences.	180
4.1	The number of training and evaluation samples used for each model tuned for correspondences of n minutiae.	192
4.2	The LR_K RMED and RMEP rates for corresponding configurations of $n = 4 \dots 9$ minutiae found in the evaluation set for the proposed model (using feature selection for dimensionality reduction) and for the method proposed in Chapter 3.	192
4.3	The LR_{weight} RMED and RMEP rates for corresponding configurations of $n = 4 \dots 9$ minutiae found in the evaluation set for the proposed model (using feature selection for dimensionality reduction) and for the method proposed in Chapter 3.	193
4.4	The LR_K rates of misleading evidence in favour of defence (RMED) and prosecution (RMEP) for corresponding configurations in the evaluation set for Experiment C in Chapter 3 and the proposed method (using subset selection and PCA for feature vector dimensionality reduction).	195
4.5	The LR_{weight} rates of misleading evidence in favour of defence (RMED) and prosecution (RMEP) for corresponding configurations in the evaluation set for experiment C in Chapter 3 and the proposed method (using subset selection and PCA for feature vector dimensionality reduction).	196
4.6	The $RMEP_R$ and $RMED_R$ of each model (for 4, 5, and 6 minutiae) for the corresponding configurations in the evaluation set resulting from different gallery sizes.	199
4.7	The $RMEP_R$ and $RMED_R$ of each model (for 7, 8, and 9 minutiae) for the corresponding configurations in the evaluation set resulting from different gallery sizes.	200
4.8	The LR_K rates of misleading evidence in favour of defence (RMED) for sub-sampled corresponding configurations of $n = 4 \dots 9$ minutiae found in the NIST27 validation set (Dataset F) containing only true correspondences.	201

4.9	The LR_K rates of misleading evidence in favour of prosecution (RMEP) for sub-sampled corresponding configurations in the Dataset G close non-match validation set containing only true correspondences.	201
5.1	The number of POI models built for each search configuration evaluated for match and close non-match crime marks.	213
5.2	The rates of misleading evidence in favour of defence (RMED) and prosecution (RMEP) for corresponding configurations of $n = 4 \dots 9$ minutiae evaluated for 1, 2, and 5 crime marks.	215

List of Publications and Presentations

Peer-Reviewed Papers

1. Abraham, J., Champod, C., Lennard, C. and Roux, C. (2013). Spatial analysis of corresponding fingerprint features from match and close non-match populations. *Forensic Sci. Int. vol. 230 no. 13*, pp. 87-98.
2. Abraham, J., Champod, C., Lennard, C. and Roux, C. (2013). Modern statistical models for forensic fingerprint examinations: A critical review. *Forensic Sci. Int. vol. 232 no. 13*, pp. 131-150.

Book Chapters

1. Abraham, J., Champod, C., Lennard, C. and Roux, C. (2013). An AFIS Candidate List Centric Fingerprint Likelihood Ratio Model based on Morphometric and Spatial Analyses (MSA), *New Trends and Developments in Biometrics, Jucheng Yang and Shan Juan Xie (eds.), ISBN: 978-953-51-0859-7, InTech*.

Conference Presentations

1. A Practical Statistical Model for Fingerprint Comparisons. *The 23rd International Symposium on the Forensic Sciences (ANZFSS 2016), Auckland, 2016*.
2. Calculating Likelihood Ratios for Fingerprint Identification ‘Cold hit’ and ‘Warm Hit’ Cases. *The International Fingerprint Research Group Conference, Patiala, India, 2015*.
3. Modelling the Variability of Minutiae using Machine Learning and Statistical Analysis to Calculate Likelihood Ratios for ‘Warm Hit’ Cases. *The 22nd International Symposium on the Forensic Sciences (ANZFSS 2014), Adelaide, 2014*.
4. Spatial Analysis of Corresponding Fingerprint Features from Match and Close Non-Match Populations. *The 6th European Academy of Forensic Science Conference, The Hague, 2012*.

5. Spatial Analysis of Corresponding Fingerprint Features from Match and Close Non-Match Populations. *The 21st International Symposium on the Forensic Sciences (ANZFSS 2012), Hobart, 2012.*

Abstract

For the majority of the 20th century, the forensic practice of fingerprint identification has had unanimous acceptance as reliable, robust, and admissible evidence. However, a number of forensic commentators have questioned the scientific validity of the current practice of fingerprint identification. Moreover, recent well publicised misidentifications have added concerns with the accuracy and quality assurance processes in practice, while fingerprint practitioners have experienced growing pressure to perform identifications from increasing workload and difficult casework.

The application of statistical modelling for fingerprint identification is a scientific methodology that provides a quantification of fingerprint evidence that can alleviate such concerns regarding the scientific foundations of fingerprint identification. Moreover, such statistical models can be used as a supportive tool for fingerprint practitioners who are under operational pressure to accurately assess crime marks against other fingermarks in a timely manner.

In this dissertation, two statistical modelling frameworks for different fingerprint identification scenarios are proposed. The first variant is called AFIS-centric models that calculate likelihood ratios and are designed to work with AFIS candidate lists, helping the practitioner to decide between match and close non-match correspondences. Two likelihood ratio measures are proposed, one with the aim of evaluating candidate list members as match or a close non-match, the other providing a weight-of-evidence evaluation.

The second model variant called a Person-of-Interest (POI) model is designed for the scenario where a rich collection of fingermarks from the same source finger are available to provide a more thorough evidential assessment. Tailored models of skin distortion are built using samples of the POI's finger, using feature vectors that make use of all of the available spatial information, from which a weight-of-evidence likelihood ratio measure is derived.

Experimental results illustrate the effectiveness of the AFIS-centric and POI models as supportive tools for casework. The significance of these research results is threefold. Firstly, the proposed AFIS-centric models illustrate how feature vector based models can focus on match and close non-match populations to provide a statistical measure agnostic of an AFIS scores that can be used for workload reduction purposes through candidate list filtering/reordering and quality assurance within the Analysis-Comparison-Evaluation-Verification (ACE-V) framework. Secondly, the proposed feature vectors add robustness and spatial completeness to the model, resulting in highly accurate models that assess real-world case samples accurately. Lastly, both proposed model variants provide a highly robust and accurate quantitative output in the form of a weight-of-evidence measure that can be used to support expert testimony.